

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 07:04:17 ; Search time 4.6765 Seconds
(without alignments)
1193.323 Million cell updates/sec

Title: US-09-939-537-35

Perfect score: 288
Sequence: 1 PRASALPAPPTGSAALPDPTQT.....VISFLILGLIGVACVLTART 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	98.6	240	2 A39016	T-cell surface gly
2	80	27.8	421	2 T30709	core protein homol
3	77	26.7	375	2 T08134	oleosin-like prote
4	76.5	26.6	710	2 D96728	hypothetical prote
5	75	26.0	539	2 T28770	hypothetical prote
6	75	26.0	891	2 G84693	probable prolina-r
7	74	25.7	464	2 S22697	extensin - Volvox
8	73	25.3	215	2 S46966	microfilariar shea
9	73	25.3	418	2 G02953	beta-3-adrenergic
10	73	25.3	482	2 T36045	hypothetical prote
11	73	25.3	614	2 E86194	hypothetical prote
12	71	24.7	262	2 T06600	acetyl-coa carboxy
13	70.5	24.5	559	2 C75286	hypothetical prote
14	70.5	24.5	660	1 Q08B3	BHLF1 protein - hu
15	69.5	24.1	1249	2 T14150	vesicle associated
16	69	24.0	279	2 T05421	hypothetical prote
17	69	24.0	405	2 S65459	beta-3-adrenergic
18	69	24.0	1151	2 T18535	high molecular mas
19	68.5	23.8	192	2 JCS876	early light-induct
20	68.5	23.8	347	2 G83171	hypothetical prote
21	68.5	23.8	884	2 T25603	hypothetical prote
22	68	23.6	801	2 T29018	hypothetical prote
23	68	23.6	907	2 E96636	hypothetical prote
24	68	23.6	1173	2 T31421	C-terminal domain-
25	68	23.6	2240	2 T37057	probable multi-dom
26	67.5	23.4	211	2 T25911	hypothetical prote
27	67.5	23.4	237	2 S46964	microfilariar shea
28	67.5	23.4	377	2 A48018	mucin 7 precursor,
29	67	23.3	191	2 F84522	probable prolina-r

30	67	23.3	272	2 148700	gene ox40 protein
31	67	23.3	318	2 A29636	myoblast determina
32	67	23.3	346	2 S19129	proline-rich prote
33	67	23.3	347	2 S10571	mucin 1 precursor,
34	67	23.3	367	1 TV0059	rod shape-determ
35	67	23.3	590	1 T35297	probable dihydrol
36	67	23.3	1119	2 T50995	related to cytoske
37	66.5	23.1	202	2 AB2378	hypothetical prote
38	66.5	23.1	350	2 E75341	peptidyl-prolyl ci
39	66.5	23.1	365	2 B42832	factor VIII intron
40	66.5	23.1	497	2 F83634	hypothetical prote
41	66.5	23.1	518	2 F70831	probable PPE prote
42	66	22.9	127	2 C85091	arabinogalactan-pr
43	66	22.9	356	2 A96826	T8K14.10 imported
44	66	22.9	756	2 T00367	hypothetical prote
45	66	22.9	1108	2 T35827	probable membrane

ALIGNMENTS

RESULT 1
A39016
T-cell surface glycoprotein CD7 precursor - human
N:Alternate names: T-cell leukemia antigen
C:Species: Homo sapiens (man)
C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C/Accession: A39016; S03520
R:Schander, L.E.; Fleenor, D.E.; Kurtzberg, J.; Haynes, B.F.; Kaufman, R.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 603-607, 1991
A>Title: Isolation and characterization of the genomic human CD7 gene: structural similar
A/Reference number: A39016, NCID:91110576; PMID:1703303
A/Accession: A39016
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <SCH>
A/Cross-references: UNIPROT:P09564; GB:M37271; NCID:G180163; PIDN:AA51953.1; PID:G180164
EMBO J 6, 3313-3316, 1987
A>Title: Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expre
A/Reference number: S03520, NCID:88111517; PMID:3501369
A/Accession: S03520
A/Molecule type: mRNA
A/Residues: 1-240 <ARU>
A/Cross-references: EMBL:X06180; NCID:G29819; PIDN:CAA29546.1; PID:G732757
C/Genetics:
A/Gene: GDB:CD7
A/Cross-references: GDB:119770; OMIM:186820
A/Map position: 17q25.2-17q25.3
A/Introns: 28/1
C/Keywords: glycoprotein; membrane protein; surface antigen; tandem repeat
F/1-25/Domain: signal sequence #status predicted <Sig>
F/26-240/Product: T-cell surface glycoprotein CD7 #status predicted <MAT>
F/145-180/Region: 9-residue repeats (X-P-P-X-A-S-A-L-P)

Query Match 98.6%; Score 284; DB 2; Length 240;
Best Local Similarity 98.3%; Pred. No. 2.3e-19;
Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRASALPAPPTGSAALPDPTQTASALPAPASALPALAAVISFLILGLIGVACVLTART 58
Db 147 PRASALPAPPTGSAALPDPTQTASALPAPASALPALAAVISFLILGLIGVACVLTART 204

RESULT 2
T30709
core protein homolog 107L - Molluscum contagiosum virus 1
N:Alternate names: MC107L
C:Species: Molluscum contagiosum virus 1
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C/Accession: T30709
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996

A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30709
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-421 <SEN>
A:Cross-references: UNIPROT:Q98274; EMBL:U60315; PIDN:AACS5235.1
C:Genetics:
C:Note: MC107L

[illegible]

RESULT 3
T08134
oleosin-like protein - rape
C/Species: Brassica napus (rape)
C/Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C/Accession: T08134
R/Hong, H.P.; Ross, J.H.; Gerster, J.L.; Rigas, S.; Datta, R.S.; Hatzopoulos, P.; Scollee
Plant Mol. Biol. 34, 549-555, 1997
A/Title: Promoter sequences from two different Brassica napus tapetal oleosin-like genes
A/Reference number: Z16373; MUID:97369377; PMID:9225865
A/Accession: T08134
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-375 <HON>
A/Cross-references: UNIPROT:P93066, EMBL:Y08986, NID:g1769971, PIDN:CAA70173.1; PID:g176
C/Genetics:
C/Introns: 105/2; 184/2

	Query Match	26.7%	Score 77;	DB 2;	Length 375;
	Best Local Similarity	47.4%;	Pred. No. 4.1;		
	Matches	18;	Conservative	1;	Mismatches 19; Indels 0; Gaps 0.
OY	1 PRASALPAEPTGSAIPDEQTASALPDPPAASLPAALA	38			
ob	213 PEAPAAPAAAPAAAPAAPAAAPAEEDDAAPAAEPAA	250			

RESULT 4
D96728
hypothetical protein F24U13.3 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D96728
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzart, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.L.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucero, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.L.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: D96728
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1..710 <STO>
A/Cross-references: UNIPROT:Q9CAL8; GB:AE005173; NID:G6175155; PIDN:AAF04882.1; GSPDB:GN
C/Genetics:
A/Gene: F24J13.3
A/Map position: 1

	Query Match	46.6%;	Score 76.5;	DB 2;	Length 710;
	Best Local Similarity	25.0%;	Pred. No. 8.2;		
	Matches	19;	Conservative	4;	Mismatches 15; Indels 3; Gaps 1;
Oy	1 PRASALPAPPT---GSAIPDPQRTASALPPDPAAMATPAL	37			
	: : : : : :				
	: : : : : :				
Dh	12 PSADSAPPPTSSDGSAAPPPTSADAPPSPPADSSPPAL	51			

RESULT 5
T28770
hypothetical protein W03D2.1 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28770
R/Rohlfing, T.; Wohldmann, P.
submitted to the EMBL Data Library, June 1997
A/Description: The sequence of *C. elegans* coamid W03D2.
A/Reference number: Z20519
A/Accession: T28770

A: Molecule type: DNA
A: Residues: 1-539 <ROH>
A: Cross-references: UNIPROT:O8MOG9; EMBL:AF000298; PIDDN:AAc48255.1; GSPDB:GN00023; CESP:V
A: Experimental source: strain Bristol N2, clone W03D2
C: GeneticS:
A: Gene: CESP:W03D2.1
A: Map position: 4
A: Introns: 40/3, 88/3, 115/3, 146/3, 173/3
A: Superfamily: collagen alpha I(VIII) chain, complement C1q carboxyl-terminal homology

Query Match	26.0%	Score 75;	DB 2;	Length 539;
Best Local Similarity	47.1%	Pred. No. 8.7;		
Matches 16;	Conservative 2;	Mismatches 14;	Indels 2;	Gaps 1;
Oy	1	PRASALPAPPPGSALEPPDQTRASALPPDPASALP	34	
	:			
db	288	FRGSPPPPPPGS--PPPPPGGSPPPPPGASPP	319	

RESULT 6
G84693
probable proline-rich protein [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: G84693
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vandenken, S.B.; Umayam, L.; Tallon, L.
euse, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A/Reference number: A84420; MUID:20083487; PMID:1061197

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-891 <STO>
A/Cross-references: UNIPROT:Q9ZM08; GB:AE002053; MID:g3980411; PIDD:AAc95214.1; GSPDE:GNM000000000
C/Genetics:
A/Gene: At2g29210
A/Map position: 2

Query Match	26.0%	Score 75;	DB 2;	Length 891;
Best Local Similarity	42.9%	Pred. No. 14;		
Matches 15;	Conservative 3;	Mismatches 11;	Indels 6;	Gaps 1.

RESULT 7
S22697
extensin - Volvox carteri (fragment)

A:Reference number: Z15788
 A:Accession: T06600
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-262 <NIE>
 A:Cross-references: UNIPROT:Q42783; EMBL:U40666; NID:g1143318; PIDD:AA67836.1; PID:g114
 A:Experimental source: strain Resnik
 C:Genetics:
 A:Gene: accB-1
 A:Genome: nuclear
 C:Complex: acetyl-CoA carboxylase is composed of biotin carboxylase (EC 6.3.4.14), carb
 C:Function: <ACC>
 A:Description: enzyme acetyl-CoA carboxylase complex catalyzes synthesis of malonyl-CoA
 carboxylation of biotin bound to biotin-carboxyl-carrier protein with bicarbonate; carb
 A:Pathway: fatty acid biosynthesis
 C:Function: <BCC>
 A:Description: biotin carboxyl carrier chain is responsible for binding of biotin
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
 C:Keywords: biotin metabolism; chloroplast; fatty acid biosynthesis; ligase
 F:1-47/Domains: transit peptide (chloroplast) #status predicted <TNP>
 F:48-262/Product: acetyl-CoA carboxylase, biotin carboxyl carrier chain #status predicted

Query Match 24.7%; Score 71; DB 2; Length 262;
 Best Local Similarity 41.7%; Pred. No. 10;
 Matches 15; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 1 PRASALPAPPTGSALPDPTQTASALPPPPASALPAA 36
 146 PPPALPPPPVASTPAPTARATPTPTSPAPVAKSA 181

RESULT 13

C75286
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: C75286
 R:White, O.; Eiben, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; PMID:20036896; PMID:10567266
 A:Accession: C75286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-559 <WHI>
 A:Cross-references: UNIPROT:Q9RRY4; GB:AE002065; GB:AE000513; NID:g6460149; PIDD:AAF1169
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2348
 A:Map position: 1
 C:Superfamily: Deinococcus radiodurans hypothetical protein DR2348

Query Match 24.5%; Score 70.5; DB 2; Length 559;
 Best Local Similarity 35.7%; Pred. No. 23;
 Matches 20; Conservative 4; Mismatches 15; Indels 17; Gaps 2;

QY 1 PRASALPAPPTGSALPDPTQTASALPD-----PPAAS--ALPAAALAV 39
 258 PATSGAPASPTTITPPATATTDPTGSDGDLTPGVTEQMPAATEPAPAPAAV 313

RESULT 14

QOBER3
 BHLF1 protein - human herpesvirus 4 (strain B95-8)
 C:Species: human herpesvirus 4; Epstein-Barr virus
 C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
 C:Accession: A03742
 R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983

A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus.
 A:Reference number: A93065; PMID:85035713; PMID:6092825
 A:Accession: A03742
 A:Molecule type: DNA
 A:Residues: 1-660 <BAN>
 A:Cross-references: UNIPROT:P03181
 R:Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; He
 Nature 310, 207-211, 1984
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A:Reference number: A03794; PMID:84270667; PMID:6087149
 A:Contents: annotation; protein coding region
 C:Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-522
 C:Superfamily: human herpesvirus 4 BHLF1 protein

Query Match 24.5%; Score 70.5; DB 1; Length 660;
 Best Local Similarity 41.0%; Pred. No. 27;
 Matches 16; Conservative 4; Mismatches 14; Indels 5; Gaps 1;

QY 1 PRASALPAPPTGSALPD-----PQTASALPDPPASALP 34
 192 PGTAPAPGPGGGAAVPSGATPHPERGSGPADPPAARLP 230

RESULT 15

T14150
 vesicle associated protein 1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T14150
 R:Shugrue, C.A.; Peters, H.; Matcovcik, L.M.; Kolen, E.R.; Czernik, A.J.; Hubbard, A.; Go
 submitted to the EMBL Data Library, November 1997
 A:Reference number: Z17887
 A:Accession: T14150
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1249 <SHU>
 A:Cross-references: UNIPROT:Q9Z2Q1; EMBL:AF034582; NID:g4104320; PIDD:g4104321; PIDD:AA020
 A:Experimental source: liver
 C:Genetics:
 A:Gene: VAP1

Query Match 24.1%; Score 69.5; DB 2; Length 1249;
 Best Local Similarity 37.3%; Pred. No. 61;
 Matches 19; Conservative 4; Mismatches 13; Indels 15; Gaps 2;

QY 1 PRASALPAPPT-----GSALPDPTQTASALP-----DPPAASALPAA 36
 965 PTSSAPLPPPPSSGASFGHGFAGAPSSSAVALPFGTTGTTPAASLELPAS 1015

Search completed: March 7, 2005, 07:21:01
 Job time : 6.8765 secs